X.0 INSTRUCTIONS FOR USING PS-11 CORRELATION TEST SPREADSHEET

This section provides instructions for using the PS-11 Correlation Test Spreadsheet. The spreadsheet is a Microsoft Excel™ file that performs the regression analyses for determining correlation equations for the five models described in Section 12.3 of PS-11. The spreadsheet also determines which model provides the best fit while satisfying all of the acceptance criteria specified in Section 13.2 of PS-11. The spreadsheet can handle up to 100 test runs on single sampling train data or up to 50 test runs on paired sampling train data.

Note: The current version of the spreadsheet (Version 2-6) accounts for several changes to PS-11 that will be incorporated into amendments to the January 12, 2004, publication of the final rule. These changes include a revised procedure for calculating the confidence and tolerance interval half ranges for the exponential and power correlation models, and revised values for one of the statistical parameters (u_{tN}) presented in Table 1 of the rule.

X.1 DESCRIPTION OF SPREADSHEET

X.1.1 Overview of the Worksheets

The spreadsheet contains several worksheets, which are listed in Table X-1 along with their functions. The worksheets appear as tabs at the bottom of the Excel screen; click on these tabs to move from one worksheet to another. Because of the number of worksheets, not all tabs are visible at one time, but you can scroll through the tabs by clicking on the arrow icons in the lower left corner of the Excel screen.

The initial worksheet ("Test Data") is used for data entry. The next 10 worksheets consist of 5 pairs of worksheets. Each pair applies to one of the five correlation models described in Section 12.3 of PS-11: linear, polynomial, logarithmic, exponential, and power. For each of these models, the first worksheet (e.g., "Linear Calcs") displays the results of the calculations for each data point. The second worksheet in the pair (e.g., "Linear Summary") displays a summary of the calculations, the correlation equation, and a summary of how the model compares to the acceptance criteria. Following the 10 model-specific worksheets is the "Best Model" worksheet, which displays how each of the models satisfies the acceptance criteria of PS-11 and identifies the model that provides the best fit for the data. The "Predicted Values" worksheet displays the PM concentrations predicted by each of the five correlation models. The "Stats Table" worksheet is a lookup table that provides the statistical parameters needed to determine the confidence and tolerance interval half range percentages. The final worksheet ("Version") indicates the spreadsheet version and date.

Table X-1. Description of Worksheets in PS-11 Correlation Test Calculation Spreadsheet

Worksheet	Function
Test Data	• Correlation test data entry sheet • Displays graph of data
Linear Calcs	• Performs the calculations for fitting a linear regression model to the correlation test data
Linear Summary	• Presents the results of the linear correlation • Includes a summary of the acceptance criteria • Displays graph of data and correlation equation
Polynomial Calcs	• Performs the calculations for fitting a polynomial regression model to the correlation test data
Polynomial Summary	• Presents the results of the polynomial correlation • Includes a summary of the acceptance criteria • Displays graph of data and correlation equation
Logarithmic Calcs	• Performs the calculations for fitting a logarithmic regression model to the correlation test data
Logarithmic Summary	• Presents the results of the logarithmic correlation • Includes a summary of the acceptance criteria • Displays graph of data and correlation equation
Exponential Calcs	• Performs the calculations for fitting an exponential regression model to the correlation test data
Exponential Summary	• Presents the results of the exponential correlation • Includes a summary of the acceptance criteria • Displays graph of data and correlation equation
Power Calcs	• Performs the calculations for fitting a power regression model to the correlation test data
Power Summary	• Presents the results of the power correlation • Includes a summary of the acceptance criteria • Displays graph of data and correlation equation
Best Model	• For each of the five models, displays - Correlation coefficient - Confidence interval half range percentage - Tolerance interval half range percentage • For polynomial model, also indicates if the maximum or minimum value occurs within the allowable range • Indicates which model provides the best fit based on correlation coefficient
Predicted Values	• For user-entered PM CEMS response values, displays the predicted PM concentrations for each of the five models
Stats Table	• Provides a table of statistical values used in determining t_f, v_f, and $u_{n'}$,
Version	• Displays the version and date of the spreadsheet

X.1.2 Locked Cells

All but two of the worksheets are completely locked, which means that you can open the worksheet, view the results of the calculations, and print out the results, but you cannot modify any of the cells in the worksheet.

The "Test Data" worksheet is partially locked. Figure X-1 shows the Test Data worksheet for an example problem. You can enter the test run numbers from the correlation test in cells A6 through A105 and the correlation test data into the block of cells between B6 and C105: the PM CEMS response data go in column B, and the PM concentrations, as determined by the appropriate reference method, go in column C. You can also enter the facility name, location, emission unit name, and test dates in cells F3 to F6, respectively, and the emission limit and emission limit units in cells F9 and G9, respectively. All other cells in the "Test Data" worksheet are locked.

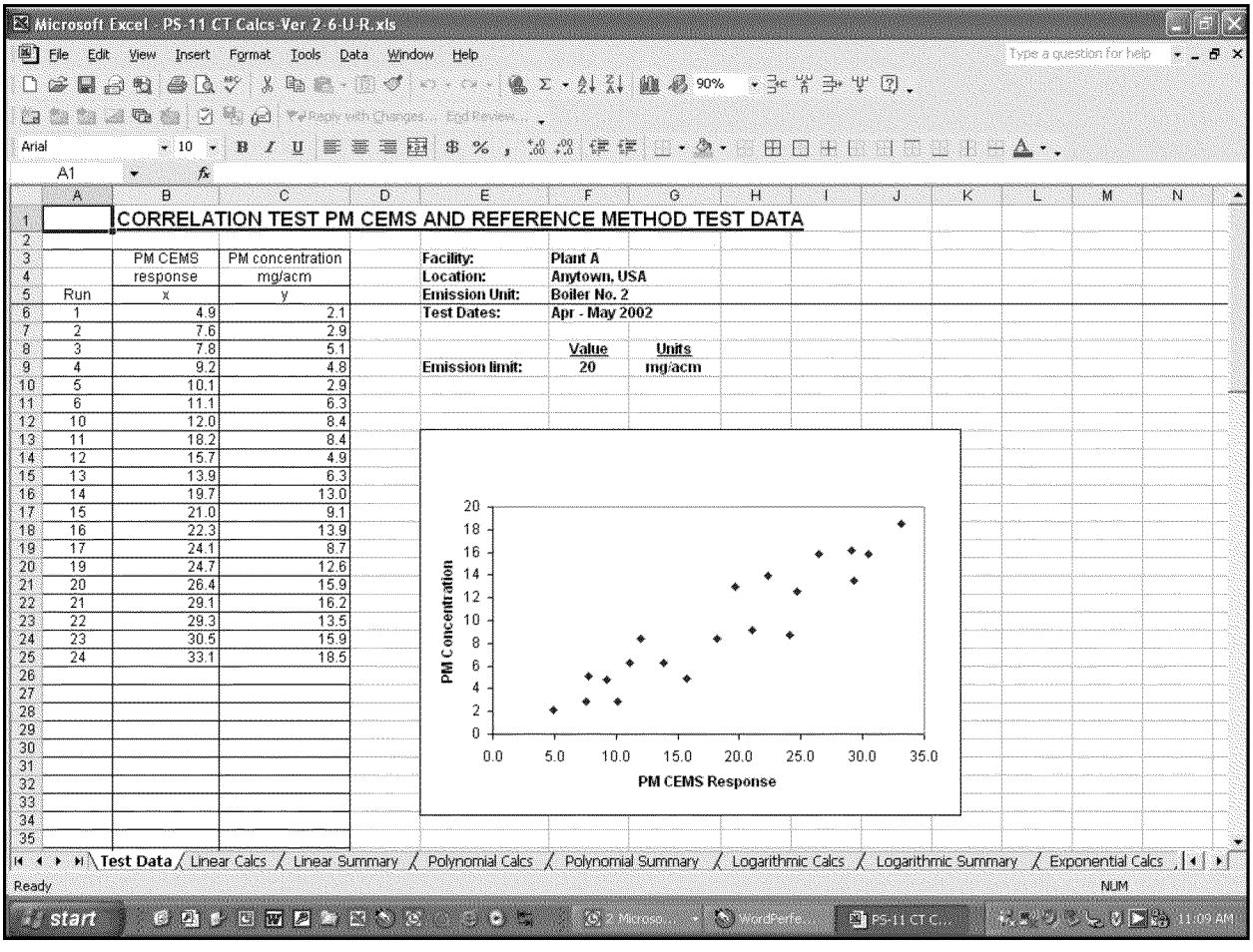
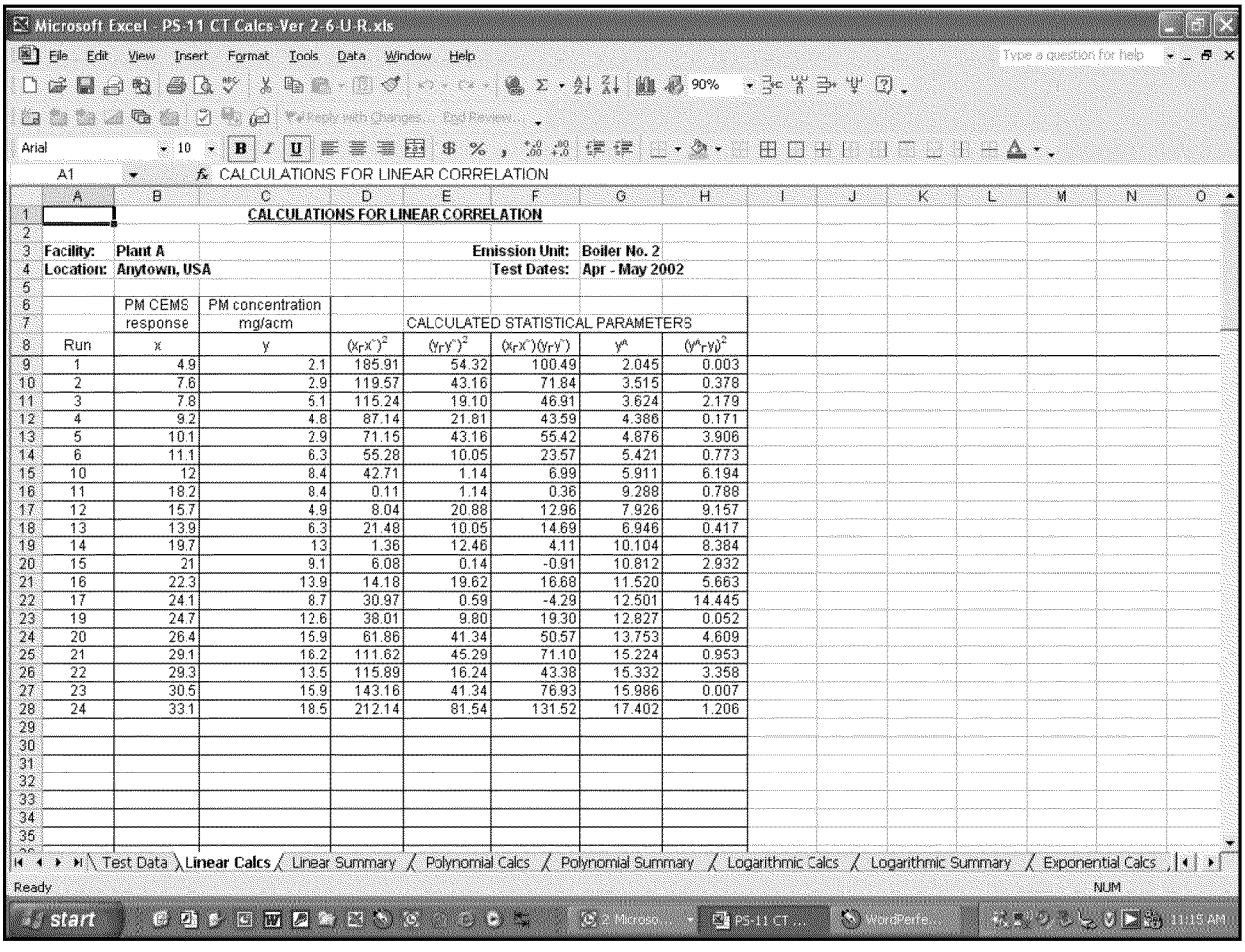


Figure X-1. Example of "Test Data" worksheet.

The "Predicted Values" worksheet is also partially locked. You can enter PM CEMS response values in cells A8 through A39 and see the resulting PM concentrations predicted by each of the correlation models. All other cells in this worksheet are locked.

X.1.3 Calculation Worksheets

For each of the five correlation models described in Section 12.3 of PS-11, the spreadsheet includes a calculation worksheet. Figure X-2 shows an example of the calculation worksheet for the linear model.



Microsoft Excel - PS-11 CT Calcs-Ver 2-6-U-R.xls

CALCULATIONS FOR LINEAR CORRELATION

Facility:	Plant A				Emission Unit:	Boiler No. 2		
Location:	Anytown, USA				Test Dates:	Apr - May 2002		
	PM CEMS response	PM concentration mg/acm	CALCULATED STATISTICAL PARAMETERS					
Run	x	y	$(x_r x')^2$	$(y_r y')^2$	$(x_r x')(y_r y')$	y^a	$(y^a r y)^2$	
1	4.9	2.1	185.91	54.32	100.49	2.045	0.003	
2	7.6	2.9	119.57	43.16	71.84	3.515	0.378	
3	7.8	5.1	115.24	19.10	46.91	3.624	2.179	
4	9.2	4.8	87.14	21.81	43.59	4.386	0.171	
5	10.1	2.9	71.15	43.16	55.42	4.876	3.906	
6	11.1	6.3	55.28	10.05	23.57	5.421	0.773	
10	12	8.4	42.71	1.14	6.99	5.911	6.194	
11	18.2	8.4	0.11	1.14	0.36	9.288	0.788	
12	15.7	4.9	8.04	20.88	12.96	7.926	9.157	
13	13.9	6.3	21.48	10.05	14.69	6.946	0.417	
14	19.7	13	1.36	12.46	4.11	10.104	8.384	
15	21	9.1	6.08	0.14	-0.91	10.812	2.932	
16	22.3	13.9	14.18	19.62	16.68	11.520	5.663	
17	24.1	8.7	30.97	0.59	-4.29	12.501	14.445	
19	24.7	12.6	38.01	9.80	19.30	12.827	0.052	
20	26.4	15.9	61.86	41.34	50.57	13.753	4.609	
21	29.1	16.2	111.62	45.29	71.10	15.224	0.953	
22	29.3	13.5	115.89	16.24	43.38	15.332	3.358	
23	30.5	15.9	143.16	41.34	76.93	15.986	0.007	
24	33.1	18.5	212.14	81.54	131.52	17.402	1.206	

Test Data | Linear Calcs | Linear Summary | Polynomial Calcs | Polynomial Summary | Logarithmic Calcs | Logarithmic Summary | Exponential Calcs

Ready NUM

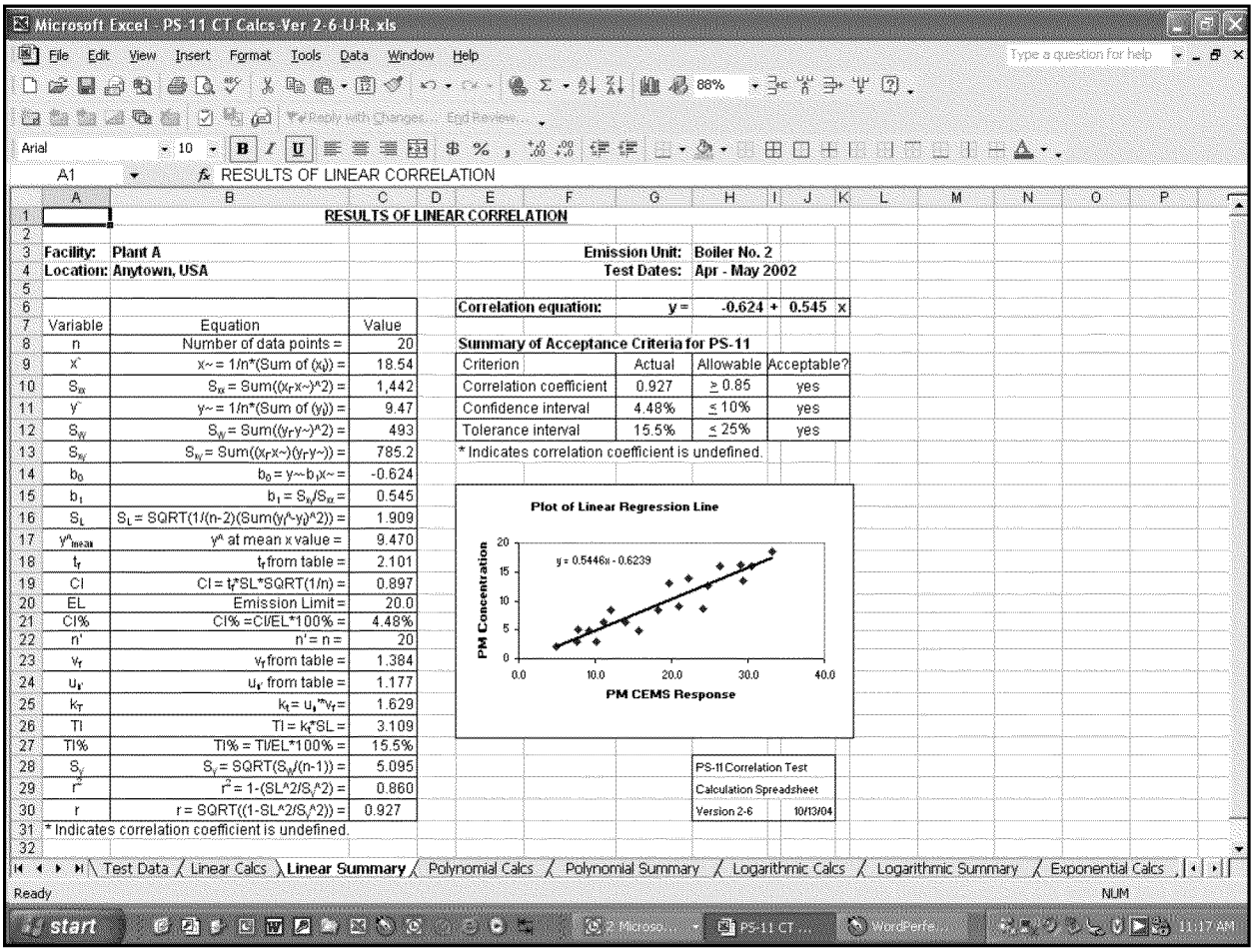
Figure X-2. Example of "Linear Calcs" worksheet.

The general format of the calculation worksheets is similar for all of the correlation models. The test run numbers, PM CEMS response (x) values, and PM concentration (y) values appear in the first columns. These values are pulled directly from the "Test Data" worksheet and do not have to be reentered here. Test run numbers appear in column A. PM CEMS response values appear in column B. If the correlation model requires the PM CEMS values to be transformed, as is the case for the logarithmic and power models, the transformed PM CEMS values appear in column C. PM concentration values appear in column C for the linear, polynomial, and exponential models. For the logarithmic and power models, the PM concentrations are displayed

in column D. The next series of columns in each "Calcs" worksheet presents the results of various intermediate calculations that are needed to determine the correlation equation. These calculations vary from model to model and are defined in the equations presented in Section 12.3 of PS-11. The last columns show the results of the calculations for the confidence interval half range, tolerance interval half range, and the other statistical parameters used to determine the confidence and tolerance interval half ranges.

X.1.4. Summary Worksheets

The summary worksheets display the equations and the results of the intermediate calculations that are used to determine the correlation equation, correlation coefficient, and confidence and tolerance interval half range percentages at the appropriate PM CEMS response value. Figure X-3 shows an example of the summary worksheet for the linear model. The first three columns show the variables, equations, and calculated values, respectively. The summary worksheets also show a summary of the acceptance criteria specified in Section 13.2 of PS-11 and indicate if the acceptance criteria were satisfied. Finally, the summary worksheets show a plot of the correlation test data and the correlation curve.



Variable	Equation	Value
n	Number of data points =	20
$\bar{x}$	$\bar{x} = 1/n*(\text{Sum of } (x_i)) =$	18.54
S_{xx}	$S_{xx} = \text{Sum}((x_i-\bar{x})^2) =$	1,442
$\bar{y}$	$\bar{y} = 1/n*(\text{Sum of } (y_i)) =$	9.47
S_{yy}	$S_{yy} = \text{Sum}((y_i-\bar{y})^2) =$	493
S_{xy}	$S_{xy} = \text{Sum}((x_i-\bar{x})(y_i-\bar{y})) =$	785.2
b_0	$b_0 = \bar{y} - b_1\bar{x} =$	-0.624
b_1	$b_1 = S_{xy}/S_{xx} =$	0.545
S_L	$S_L = \text{SQRT}(1/(n-2)(\text{Sum}(y_i-\hat{y}_i)^2)) =$	1.909
$\hat{y}_{mean}$	$\hat{y}$ at mean x value =	9.470
t_f	t_f from table =	2.101
CI	$CI = t_f*SL*\text{SQRT}(1/n) =$	0.897
EL	Emission Limit =	20.0
CI%	CI% = CI/EL*100% =	4.48%
n'	n' = n =	20
v_f	v_f from table =	1.384
$u_{f'}$	$u_{f'}$ from table =	1.177
k_T	$k_T = u_{f'}*v_f =$	1.629
TI	$TI = k_T*SL =$	3.109
TI%	TI% = TI/EL*100% =	15.5%
S_y	$S_y = \text{SQRT}(S_{yy}/(n-1)) =$	5.095
r^2	$r^2 = 1-(SL^2/S_y^2) =$	0.860
r	$r = \text{SQRT}((1-SL^2/S_y^2)) =$	0.927

* Indicates correlation coefficient is undefined.

Figure X-3. Example of "Linear Summary" worksheet.

X.1.5 Best Model Worksheet

This worksheet provides a summary of the results for all of the models. Figure X-4 shows an example of the "Best Model" worksheet. For each model, the correlation coefficient, confidence interval half range percentage, and tolerance interval half range percentage are listed with an indication of whether the model meets the acceptance criterion for each of these parameters. For the polynomial model, the worksheet indicates if the model satisfies the criterion for minimum or maximum value specified in Section 12.4(3) of PS-11. The worksheet shows whether each model satisfies all of the criteria for PS-11. The worksheet identifies the best model based on the correlation coefficients for the models that meet all of the acceptance criteria.

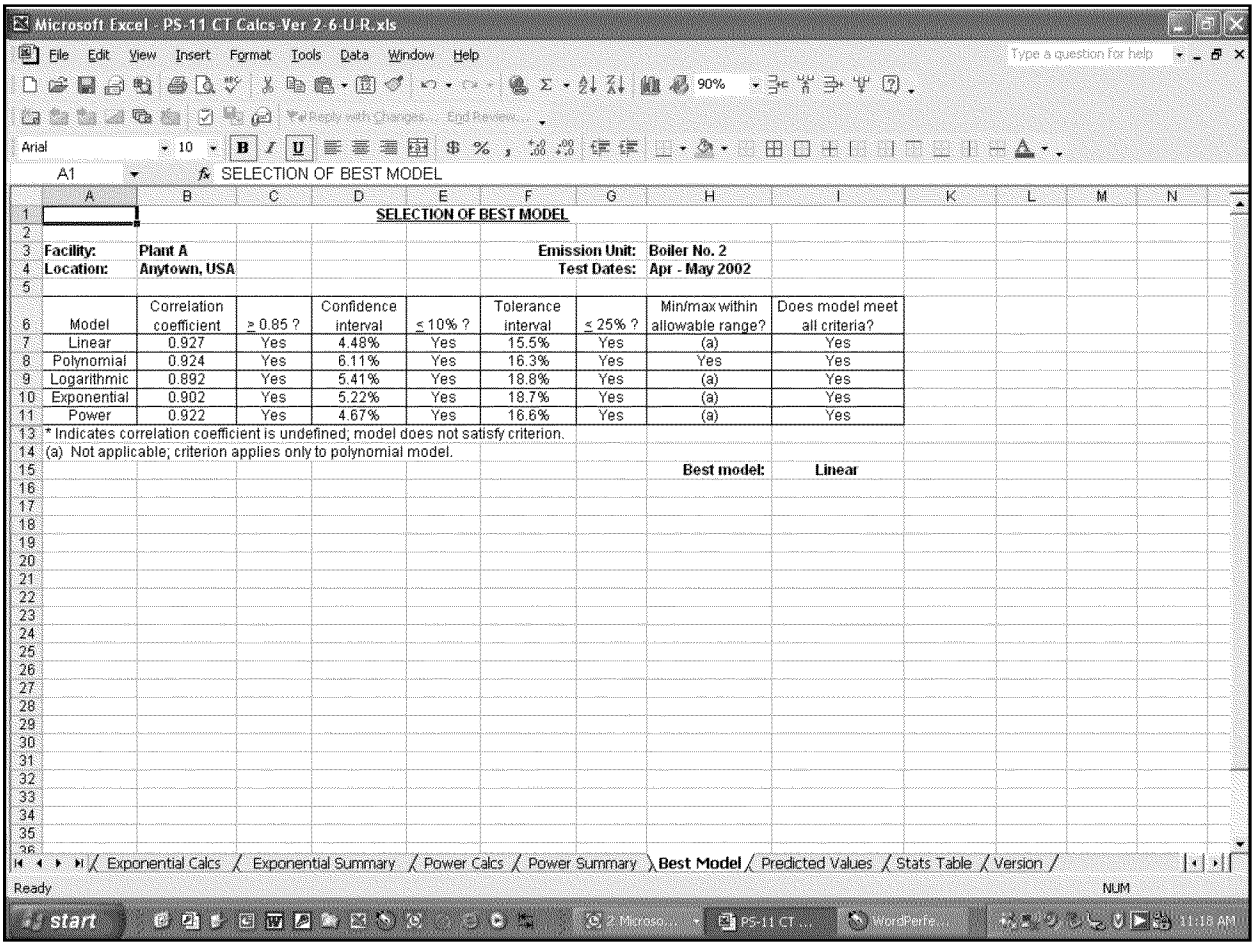
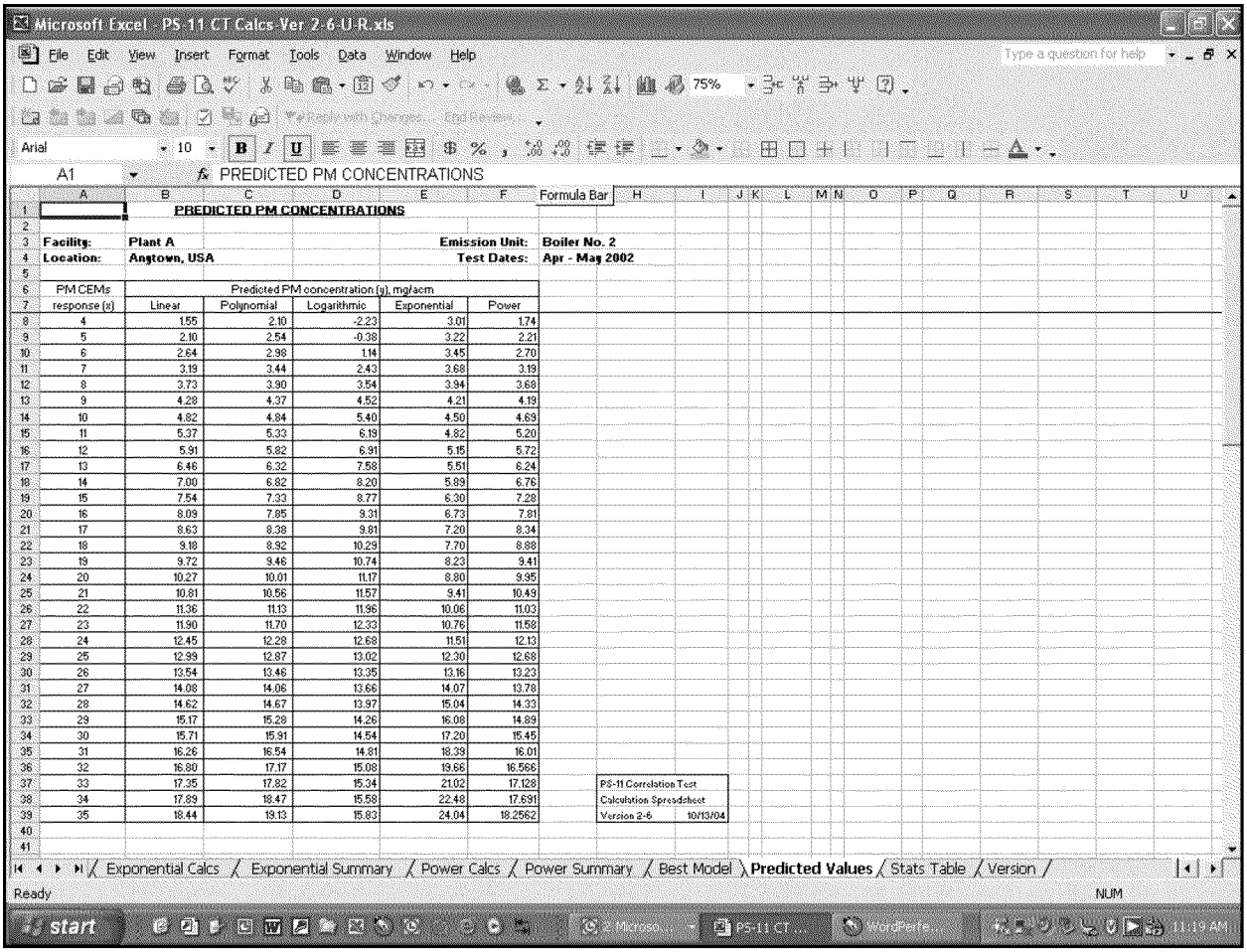


Figure X-4. Example of "Best Model" worksheet.

X 1.6 Predicted Values Worksheet

This worksheet displays the predicted concentrations for each of the five models as a function of any PM CEMS response values that are entered in cells A8 to A39. Figure X-5 shows an example of the "Predicted Values" worksheet.



PM CEMs response (x)	Predicted PM concentration (y), mg/acm Linear	Polynomial	Logarithmic	Exponential	Power
4	1.55	2.10	-2.23	3.01	1.74
5	2.10	2.54	-0.38	3.22	2.21
6	2.64	2.98	1.14	3.45	2.70
7	3.19	3.44	2.43	3.68	3.19
8	3.73	3.90	3.54	3.94	3.68
9	4.28	4.37	4.52	4.21	4.19
10	4.82	4.84	5.40	4.50	4.69
11	5.37	5.33	6.19	4.82	5.20
12	5.91	5.82	6.91	5.15	5.72
13	6.46	6.32	7.58	5.51	6.24
14	7.00	6.82	8.20	5.89	6.76
15	7.54	7.33	8.77	6.30	7.28
16	8.09	7.85	9.31	6.73	7.81
17	8.63	8.38	9.81	7.20	8.34
18	9.18	8.92	10.29	7.70	8.88
19	9.72	9.46	10.74	8.23	9.41
20	10.27	10.01	11.17	8.80	9.95
21	10.81	10.56	11.57	9.41	10.49
22	11.36	11.13	11.96	10.06	11.03
23	11.90	11.70	12.33	10.76	11.58
24	12.45	12.28	12.68	11.51	12.13
25	12.99	12.87	13.02	12.30	12.68
26	13.54	13.46	13.35	13.16	13.23
27	14.08	14.06	13.66	14.07	13.78
28	14.62	14.67	13.97	15.04	14.33
29	15.17	15.28	14.26	16.08	14.89
30	15.71	15.91	14.54	17.20	15.45
31	16.26	16.54	14.81	18.39	16.01
32	16.80	17.17	15.08	19.66	16.566
33	17.35	17.82	15.34	21.02	17.128
34	17.89	18.47	15.58	22.48	17.691
35	18.44	19.13	15.83	24.04	18.2562

Figure X-5. Example of "Predicted Values" worksheet.

X.2 HOW TO USE THE SPREADSHEET

X.2.1 Data Entry

Using the spreadsheet entails entering the correlation test data and the emission limit into the "Test Data" worksheet. The spreadsheet is designed to handle up to 100 data points and can be modified to handle more. As you enter data, the data are displayed on a graph that appears on the same worksheet.

As indicated in Section 8.6(1)(iii) of PS-11, reference method test results must first be converted to units of mass concentration that are consistent with the conditions measured by the PM CEMS. For example, if the PM CEMS responds in terms of actual conditions, the PM emission data should be converted to units of milligrams per actual cubic meter (mg/acm). You should perform this data conversion before you enter the data into the spreadsheet.

For each test run, enter the run number, PM CEMS response value, and the PM concentration into columns A, B, and C, respectively. You can enter the data in any order; for example, it is not necessary to order the data by concentration from lowest to highest. If the PM data were measured using paired sampling trains, each PM concentration should be entered as a separate data point (x value), as specified in Section 12.3 of PS-11. In such cases, you would enter the same PM CEMS response value twice, once for each of the two sampling train data points. For example, assume that the PM concentrations measured by Trains A and B during Run 3 were 6.8 and 7.3 mg/acm, respectively, and the PM CEMS response for Run 3 was 5.4 mA. For Run 3A, you would enter 6.8 mg/acm as the x value and 5.4 mA as the y value. For Run 3B, you would enter 7.3 mg/acm as the x value and 5.4 mA as the y value.

To determine the confidence interval and tolerance interval half range percentages, you must also enter the emission limit value and units (e.g., mg/acm, mg/dscm) into cells F9 and G9, respectively.

X.2.2 Checking Results

Once you have entered the test data, the spreadsheet automatically calculates the correlation equation, correlation coefficient, confidence interval half range percentage, and tolerance interval half range percentage for each of the five correlation models. A summary of the results is displayed on the "Best Model" worksheet, as shown in Figure X-4 for an example data set. In this example, all five models meet all of the acceptance criteria. The linear model has the highest correlation coefficient, with an r value of 0.927. The confidence interval half range percentage at the mean PM CEMS response value is 4.48 percent, which meets the criterion of less than or equal to 10 percent. The tolerance interval half range percentage at the mean PM CEMS response value is 15.5 percent, which meets the acceptance criterion of less than or equal to 25 percent.

Note: For data that show little to no correlation, the correlation coefficient (r) is undefined because the equation used to calculate r requires taking the square root of a negative number. In such cases, the spreadsheet displays an asterisk instead of a number for the correlation coefficient, and indicates that the data do not meet the correlation coefficient criterion of $0.85 specified in Section 13.2 of PS-11.

You can see more detailed results for any of the correlation models by clicking on the appropriate worksheet tab (e.g., Linear Summary, Polynomial Summary). As shown in Figure X-3, the value of each statistical parameter used in the calculations for the selected model is displayed on the left side of the worksheet. On the right side of the worksheet is a summary of the acceptance criteria specified in Section 13.2 of PS-11, the values for the specific data set and models, and an indication of whether or not the model satisfies those criteria. For all models, the parameters displayed in the summary are the correlation coefficient, confidence interval half range as a percentage of the emission limit, and the tolerance interval half range as a percentage of the emission limit. The correlation curve and equation also are displayed on the data plot in the summary worksheets.

Section 12.4(3) of PS-11 specifies additional criteria for minimum and maximum values for polynomial correlations. Each polynomial correlation curve has a minimum or maximum value,

which must occur outside the allowable range of values. The minimum allowable value corresponds to the minimum PM CEMS response value used to develop the correlation equation. The maximum allowable value corresponds to 1.25 times the maximum PM CEMS response value used to develop the correlation equation. Therefore, the summary of acceptance criteria displayed on the "Polynomial Summary" worksheet also indicates whether the correlation equation satisfies these criteria for minimum and maximum PM CEMS response values.

X.2.4 Predicted Values

The "Predicted Values" worksheet displays the PM concentrations predicted by each of the five correlation models for a range of PM CEMS response values. You can enter any PM CEMS response value into the appropriate cells in column A of the worksheet, and the predicted concentrations are automatically calculated and displayed, as shown in Figure X-5.

X.2.5 Printing Results

To print the results of any worksheet, click on the print button at the top of the spreadsheet. The worksheet will be printed with the file name, worksheet name, and date displayed in the footer of the page.